



Clinical Laboratories Collaborate to Resolve Differences in Variant Interpretations Submitted to ClinVar

Citation

Harrison, Steven M., Jill S. Dolinsky, Amy E. Knight Johnson, Tina Pesaran, Danielle R. Azzariti, Sherri Bale, Elizabeth C. Chao, Soma Das, Lisa Vincent, and Heidi L. Rehm. 2017. "Clinical Laboratories Collaborate to Resolve Differences in Variant Interpretations Submitted to ClinVar." *Genetics in medicine : official journal of the American College of Medical Genetics* 19 (10): 1096-1104. doi:10.1038/gim.2017.14. <http://dx.doi.org/10.1038/gim.2017.14>.

Published Version

doi:10.1038/gim.2017.14

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:34492223>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)



Published in final edited form as:

Genet Med. 2017 October ; 19(10): 1096–1104. doi:10.1038/gim.2017.14.

Clinical Laboratories Collaborate to Resolve Differences in Variant Interpretations Submitted to ClinVar

Steven M. Harrison, PhD^{1,2}, Jill S. Dolinsky, MS³, Amy E. Knight Johnson, MS⁴, Tina Pesaran, MA, MS³, Danielle R. Azzariti, MS¹, Sherri Bale, PhD⁵, Elizabeth C. Chao, MD^{3,6}, Soma Das, PhD⁴, Lisa Vincent, PhD⁵, and Heidi L. Rehm, PhD^{1,2,7,8}

¹Laboratory for Molecular Medicine, Partners HealthCare Personalized Medicine, Cambridge, MA, USA

²Harvard Medical School, Boston, MA

³Ambry Genetics, Aliso Viejo, CA, USA

⁴Department of Human Genetics, Clinical Molecular Genetics Laboratory, The University of Chicago, Chicago, IL, USA

⁵GeneDx, Gaithersburg, MD, USA

⁶Department of Pediatrics, Division of Genetics and Genomics, University of California Irvine, Irvine, CA, USA

⁷Department of Pathology, Brigham & Women's Hospital, Boston, MA, USA

⁸The Broad Institute of MIT and Harvard, Cambridge, MA, USA

Abstract

Purpose—Data sharing through ClinVar offers a unique opportunity to identify interpretation differences between laboratories. As part of a ClinGen initiative, four clinical laboratories (Ambry, GeneDx, Partners Healthcare Laboratory for Molecular Medicine, and University of Chicago Genetic Services Laboratory) collaborated to identify the basis of interpretation differences and to investigate if data sharing and reassessment resolves interpretation differences by analyzing a subset of variants.

Methods—ClinVar variants with submissions from at least two of the four participating laboratories were compared. For a subset of identified differences, laboratories documented the basis for discordance, shared internal data, independently reassessed with the ACMG-AMP guidelines, and then compared interpretations.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Corresponding author: Steven Harrison, 65 Landsdowne St, Cambridge, MA 02139, Ph: (617) 768-8269; Fax: (617) 768-8513, sharrison6@partners.org.

Supplementary information is available at the *Genetics in Medicine* website.

Conflicts of Interest

All authors are clinical service providers and are employed by laboratories which offer fee-based clinical sequencing. This employment is noted in the author affiliations. All authors declare no additional conflicts of interest beyond their employment affiliation.

Results—6,169 variants in ClinVar were interpreted by at least two of the participating laboratories, of which 88.3% were initially concordant. Laboratories reassessed 242/724 initially discordant variants, of which 87.2% (211) were resolved by reassessment with current criteria and/or internal data sharing. 12.8% (31) of reassessed variants remain discordant due to differences in application of the ACMG-AMP guidelines.

Conclusion—Participating laboratories increased their overall concordance from 88.3% to 91.7%, indicating that sharing variant interpretations in ClinVar, allowing identification of differences and motivation to resolve those differences, is critical to move toward more consistent variant interpretations.

Keywords

ClinVar; variant interpretation; data sharing; ACMG-AMP guidelines

INTRODUCTION

Interpreting the clinical significance of genetic variants is a challenging process that involves gathering and assessing available evidence, followed by formal classification based on this evidence. Variant classifications from laboratories may confirm the cause of disease, illuminate an etiology not previously considered, or guide treatment decisions. Historically, variant classifications from clinical laboratories have mostly been unavailable to the larger genetics community except for the small fraction published in journals. Recent efforts by the National Institute of Health (NIH) funded Clinical Genome Resource (ClinGen; <https://www.clinicalgenome.org/>) to support widespread data sharing using the ClinVar database has accelerated the public sharing of interpreted variants, particularly by clinical laboratories.¹ The ClinVar database (<http://www.ncbi.nlm.nih.gov/clinvar/>), maintained by the National Center for Biotechnology Information (NCBI), archives and aggregates submitted interpretations of the clinical significance of variants, with opportunities for each submitter to provide supporting evidence, and indicates if the submitted interpretations are concordant or discordant.² Sharing variants in ClinVar allows for crowd-sourcing the enormous labor of variant interpretation and provides transparency and inherent “peer review” of variant classifications. Additionally, sharing data facilitates identification of differences in variant interpretation, allowing a valuable opportunity for clinical laboratories and other ClinVar submitters to resolve those differences.

Prior reports of inconsistencies in variant interpretations have ranged from 53% discordance in double review of pathogenic or likely pathogenic variants³, 53% discordance of uncertain significance interpretations from one clinical laboratory compared to another⁴, and 60% discordance of a single laboratory’s interpretations to other clinical laboratories⁵. However, these comparisons likely overestimate discordance as most compare a current interpretation to historic interpretations from other source(s), which may be outdated and not based on an assessment of the same evidence. Reported discordant counts are typically limited to one- or two-step differences between the three major classification levels: “pathogenic (P)/likely pathogenic (LP)”, “uncertain significance (VUS)”, and “likely benign (LB)/benign (B)”. Differences in confidence or “likelihood” are not included in these discordant counts as they are unlikely to impact clinical care and are not reported as conflicts in ClinVar.^{2,7}

Differences in variant interpretations between clinical laboratories often occur due to differences in algorithms for weighing evidence to reach a classification or differences in available evidence, such as internal frequency and data or access to proprietary databases. A recent report found that 33% of interpretation differences between clinical laboratories were due to lack of access to privately held data.⁵

Historically, clinical laboratories had developed their own internal methods for variant assessment and classification, and some relied heavily upon claims in the peer-reviewed literature, both contributing to inconsistencies in variant classification between laboratories. In 2015, the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) published a joint guideline for variant interpretation that provides an evidence-based framework to classify variants, requiring laboratories to critically review literature claims and base classifications on a review of the primary evidence.⁸ By including a defined set of evidence types, the ACMG-AMP guidelines enable tracking of the specific evidence types and strengths used to classify each variant, which facilitates classification rationale transparency and accurate classification comparisons. However, given the complexity of variant interpretation, application of the criteria still requires subjective interpretation. A recent study piloting the 2015 ACMG-AMP guidelines across nine sites involved in the Clinical Sequencing Exploratory Research (CSER) program⁹ found that classifications were initially 43.4% (43/99) discordant but resolution was possible for 67.4% (29/43) of initial interpretation differences leaving only 5% (5/99) of variants with medically significant differences (P/LP vs. VUS/LB/B) remaining at the conclusion of the study.⁷ This exercise, however, was performed before laboratories had operationally implemented the guidelines, and for the majority of variants, was not based upon variants encountered during routine clinical testing. In addition, in May 2015, ClinVar was found to have 17% discordance in variants with multiple submitters, but this discordance is also based on all interpretations submitted to ClinVar, which includes interpretations from research laboratories, locus-specific databases, and aggregate databases (Online Mendelian Inheritance in Man [OMIM] and GeneReviews).¹ Therefore, we sought to examine variant interpretation concordance in the routine clinical laboratory environment.

Four clinical laboratories, Ambry Genetics, GeneDx, Partners HealthCare Laboratory for Molecular Medicine (LMM), and University of Chicago Genetic Services Laboratories (UCGSL), are working together to identify interpretation differences after submission to ClinVar and collaborate to resolve these differences by data sharing and reassessment with laboratories' current criteria, consistent with ACMG-AMP guidelines. While complete concordance in interpretations is unlikely due to subjectivity in weighting evidence and interpretation of guidelines, our goal was to track detailed data on a subset of all interpretation differences to identify trends regarding the basis of interpretation differences between laboratories and investigate the ability of laboratories to resolve differences in variants already clinically reported, which will educate the community on how to move towards more consistent variant interpretations.

MATERIALS AND METHODS

ClinVar Data Analysis

A custom report was generated by ClinVar on January 6, 2016 to identify all alleles (measure_ids) in ClinVar with submissions from at least two of the four participating laboratories (Ambry, GeneDx, LMM, and UCGSL) and report the interpretations from all participating laboratories for each identified allele. For each allele, interpretations were compared to determine concordance or discordance; and if discordant, the level of discordance was assigned. For variants with greater than two interpretation terms submitted, the two most discordant terms were used to assign the level of discordance. For example, a variant with a pathogenic, likely pathogenic, and uncertain significance interpretation would be categorized by the pathogenic vs. uncertain significance difference. Metrics were calculated for all four laboratories in aggregate as well as each pair-wise comparison of the four laboratories to facilitate identification of variants for reassessment.

Resolution Process

Each laboratory compared their interpretations in ClinVar to their current internal interpretations to determine if the difference was resolved by a more recent reassessment not yet represented in ClinVar. Next, each pair of laboratories identified a subset of variants with interpretation differences for reassessment (242 variants of the total 724 variants with interpretation differences), prioritizing variants with P/LP vs. LB/B differences, followed by additional medically significant differences, and genes and/or disease areas with a higher frequency of interpretation differences. After identifying variants for reassessment, participating laboratories shared internal data, discussed initial classification rationale, and categorized the reason for the initial interpretation difference when it was apparent (e.g. differences in classifications rules or internal data). After discussion, laboratories independently reassessed variants with the ACMG-AMP guidelines and, if appropriate, reclassified the variant in their own database. Laboratories continued discussion for variants with interpretations that remained different after internal reassessment allowing consensus to be achieved for an additional set of discordant variants. For variants remaining discordant, any unique ACMG-AMP criteria applied by one laboratory but not the other participating laboratory was tracked. For statistical comparisons, a Fisher's Exact test was used.

RESULTS

Initial Concordance Data

As of January 2016, over 49,000 unique variants had been submitted to ClinVar by at least one of the four participating laboratories (Ambry, GeneDx, LMM, and UCGSL). Comparison of data in ClinVar identified 6,169 variants from 308 genes interpreted by at least two of the four participating laboratories. This large number of variants interpreted by at least two participating clinical laboratories facilitated comparison of interpretations and potential resolution of differences.

Of the 6,169 shared variants (Figure 1A), 88.3% of interpretations between laboratories were concordant, with 62.3% (3,845 variants) in exact agreement and 26.0% (1,600 variants)

differing only in confidence levels (P vs. LP and LB vs. B). As confidence differences are unlikely to impact clinical care or management, they were not considered discordant for this project, and are included in concordant counts. The remaining 11.7% (724) of shared variants had discordant interpretations and were categorized by the level of discordance. 3.5% (216 variants) of all shared variants were medically significant differences (MSDs), pathogenic (P/LP) versus other (VUS/LB/B), that are most likely to impact medical management; however, the majority of MSDs (94.4%) were P/LP versus VUS differences and only 12 variants of this set (5.6%) had P/LP versus LB/B differences. The remaining 8.2% (508 variants) of all shared variants were VUS versus benign (LB/B) differences, which are less likely to affect medical management, but may lead to differences in counseling time and medical management. Analysis of concordance and discordance by variant type showed highest concordance (97.0%; 479/494) for interpretation of predicted null variants (nonsense, frameshift, and +/-1,2 splice site) with these differences accounting for only 0.8% (4/508) of all VUS vs LB/B differences and 5.1% (11/216) of all MSDs. Interpretation of missense variants were 84.6% concordant (3,053/3,607) but, as expected, missense differences accounted for the highest percent of both VUS vs LB/B differences (70.9%; 360/508) and MSDs (89.8%; 194/216).

The 724 variants with interpretation differences were from 148 genes. To determine trends by disease area, genes were categorized into four categories (cardiovascular, hereditary cancer, neurologic, and other) and the proportion of MSDs and other differences are shown in the first column for each disease area ("Total"; Figure 2). See Supplemental Table 1 for pre and post-resolution concordance rates in each disease area. The neurologic disease area had the lowest initial concordance (81.2%) with the majority of differences being VUS vs. LB/B compared to MSDs (17.8% and 1.0% of all neurologic variants, respectively). Cardiovascular was the only disease area with more initial MSDs than VUS vs. LB/B differences (5.9% and 5.2% of all cardiovascular variants, respectively). A summary for each of the 148 genes with interpretation differences, including the number of shared variants and percent concordant versus discordant, is provided in Supplemental Table 2.

Interpretation Differences Resolution

To facilitate resolution of interpretation differences between laboratories, each pair-wise combination of the four participating laboratories were compared to determine the number of interpretation differences from each pair of laboratories (Supplemental Table 3). Laboratory pairs with a higher number of total interpretation differences selected variants for reassessment based on level of discordance (prioritizing variants with medically significant differences, especially those with P/LP vs. LB/B differences) and by disease area, with cardiovascular reassessments prioritized by genes with multiple interpretation differences (*MYBPC3*, *GLA*, *ACTC1*, *MYL2*, *MYL3*, *TPM1*, *RAF1*) and hereditary cancer reassessments prioritized by variants in high-risk hereditary breast and ovarian cancer genes or genes with National Comprehensive Cancer Network (NCCN) criteria for management (*BRCA1*, *BRCA2*, *CDH1*, *PTEN*, *TP53*). Laboratory pairs with lower total interpretation differences prioritized variants for reassessments simply based on the level of discordance. In total, laboratories reassessed 242 variants with interpretation differences (116 variants with MSDs and 126 variants with other differences) to determine if sharing internal data and

applying the ACMG-AMP criteria and classification rules could resolve the different interpretations (Figure 3).

For the 116 reassessed variants with MSDs, 79.3% (92 variants) were resolved by reassessment, data sharing and discussion with 19.8% (23 variants) resolved as P/LP, 58.6% (68 variants) resolved as VUS, and 0.9% (1 variant) resolved as LB. For the 126 reassessed variants with VUS vs. LB/B differences, 94.4% (119 variants) were resolved with 16.7% (21 variants) resolved as VUS and 77.8% (98 variants) resolved as LB/Bn. In total, 211 of 242 reassessed variants (87.2%) were resolved. Comparing disease areas (Figure 2), laboratories had a significantly higher resolution rate for cardiovascular variants than hereditary cancer variants ($p=0.00059$) and neurologic variants ($p=0.027$); no significant differences were observed between hereditary cancer and neurologic variants (Figure 2). The difference between cardiovascular and hereditary cancer resolution rates is due to a significantly higher resolution rate specifically for cardiovascular MSDs (74/77) versus hereditary cancer MSDs (17/32; $p=0.0001$). No significant differences in resolution of VUS vs LB/B interpretation differences were observed between disease areas. See Supplemental Table 4 for a full list of resolved variants. As classifications may change over time with available evidence, please check ClinVar for the most up-to-date classifications.

For a subset of variants with resolved interpretation differences (100 variants), the reason for the initial interpretation differences and whether reassessment with ACMG/AMP guidelines and/or sharing internal evidence impacted the re-classification were both documented (Figure 4). For each variant, laboratories decided which of the pre-defined categories best explained the initial difference in interpretation. Reassessment of older variant interpretations with updated classification criteria, consistent with ACMG-AMP guidelines, resolved 36% of variants (Figure 4, red shading). Sharing internal data facilitated resolution of 33% of resolved variants (blue shading), with sharing segregation data (10%) and co-occurrence data (9%) facilitating the largest percent of variants. Differences in the use or weighting of public data accounted for 14% of interpretation differences (purple shading), including benign/likely benign thresholds (9%), and different data sources (5%). Lastly, for 17% of resolved variants, the interpretation differences were found to have already been resolved but the new interpretations were not yet submitted to ClinVar.

Persistent Interpretation Differences

After reassessment and data sharing, laboratories were unable to reach consensus for 12.8% (31 variants) of the 242 reassessed variants, of which 24 are unresolved MSDs and 7 are unresolved other differences (Supplemental Table 5). For all unresolved differences, reassessments from the submitting laboratories were compared to determine the unique ACMG-AMP criteria applied that accounted for the different clinical interpretations (Table 1). Unique ACMG-AMP criteria applied were then grouped into evidence categories, as used in Richards et al 2015 Figure 1⁸, to determine if specific categories of evidence were more likely to be differently applied than others. For the persistent MSDs, 50% (12 variants) were impacted by differences in the application of benign criteria, such as: minor allele frequency data (BS1), observation in unaffected adults (BS2), and functional studies showing no deleterious effect (BS3). The remaining persistent MSDs were due to different

application of pathogenic criteria, with evidence in the functional data category impacting the highest percent of variants. For the 7 variants with unresolved VUS vs. LB/B differences (Table 1), the most frequent differentially applied ACMG-AMP criteria was minor allele frequency data (BS1). In total, across both discordant types, evidence in the functional data category (48%) and population data category (45%) contributed to the highest number of persistent interpretation differences.

Additionally, even though interpretations remained discordant, for 7 variants, reassessment reduced the level of discordance from a two-step difference (P/LP vs. B/LB) to a one-step difference (P/LP vs. VUS or VUS vs. LB/B). These variants are highlighted in grey in Supplemental Table 5.

DISCUSSION

Data sharing through ClinVar offers a unique opportunity to identify classification differences between laboratories and to collaborate to resolve these differences. By submitting variant interpretations to ClinVar and working together, four clinical laboratories were able to resolve differences in 87.2% of initially discordant variants that were reviewed (211/242 variants). With only 33.4% of all variants with interpretation differences reviewed so far (242/724 variants), participating laboratories have already increased their overall concordance rate from 88.3% to 91.7% (Figure 1B), indicating that sharing internal evidence and classification rationales is critical to move toward more consistent variant interpretations. This concordance rate is higher than previous reports because the scope was limited to a set of four clinical laboratories, as opposed to a report of 77% concordance between all ClinVar submitters of a set of variants⁶, and because participating laboratories shared all evidence such that interpretations are based on the same set of evidence, as opposed to a report of only 40% concordance between laboratories when internal data was not shared⁵. Additionally, both initial concordance (88.3%) and resolution rate (87.2%) are higher than reported by the nine sites in the CSER program (56.6% and 72.5%, respectively)⁷ as the majority of selected CSER variants were outside the scope of variants and genes encountered by the CSER sites during routine clinical testing, suggesting that expertise in disease areas may contribute to increased concordance.

Implementation of the evidence-based 2015 ACMG-AMP guidelines for sequence variant interpretation has alleviated some burden of differences due to classification algorithms. Thirty-six percent of resolved variants were initially discordant due to differences in each laboratory's internal classification method used at the time of initial assessment but the interpretation differences were resolved by reassessment with laboratories' current criteria, consistent with ACMG-AMP guidelines. However, implementation of the ACMG-AMP guidelines still requires professional judgment in deciding which evidence types are met, as 12.8% of variants reassessed with the guidelines were unable to be resolved. It is important to note that for the persistent interpretation differences, laboratories had access to the same data, including shared internal data; however, the interpretation of that data in the context of meeting specific ACMG-AMP criteria differed. The majority of persistent interpretation differences were due to differences in the application of functional data (48%) and population data (45%), both of which require gene or disease-specific knowledge.

Differences in the application of these criteria suggest that further gene specification of the ACMG-AMP guidelines by experts, such as providing specific allele frequency cutoffs and indicating which functional assays are considered well-established, could increase concordance. Specifically, comparison of the reassessment outcome between four disease areas indicates that genes in the inherited cancer disease area would greatly benefit from ACMG-AMP specification as this disease area had the lowest rate of resolution (78%) and highest rate of persistent medically significant interpretation differences (Figure 2; Supplemental Table 1). Recent publications have provided feedback and recommendations for the ACMG-AMP guidelines.^{7,10,11} ClinGen disease-specific working groups and the ClinGen Sequence Variant Interpretation working group are focused on specifying the ACMG-AMP guidelines for diseases of interest, including minor allele frequency thresholds, gene-specific functional assay and domain data, as well as providing further quantitative guidance on individual criteria such as segregation data, population data and case data. These recommendations will become public once completed and approved by ClinGen.

Our analyses demonstrated that resolution of 33% of variant interpretation differences benefited from sharing internal data. To date, such data has been typically inaccessible except for the small subset of data published in the literature, often many years after it was identified. This leads to important information that could affect patient diagnoses being unavailable to those individuals who could benefit. However, through submission to ClinVar, laboratories can more quickly share this data as it accumulates, allowing other laboratories and the patients they serve to benefit. It should be noted that some, but not all, clinical laboratories provide detailed supporting evidence in ClinVar, in the form of citations, text-based interpretation summaries, and case-level observations. Still, even for laboratories that have not yet shared this detail, ClinVar submissions at least identify that a laboratory has observed the variant and prompts further collaborative investigations and data sharing when patient treatment decisions could be impacted.

Over a third of the variant interpretation differences were resolved simply by one laboratory reassessing an old variant interpretation with the laboratory's current criteria, consistent with ACMG-AMP guidelines. Most clinical laboratories reassess variants when that variant is observed in an additional case or at the request of providers. Thus, for many remaining interpretation differences not yet reviewed, routine clinical laboratory reassessment with current criteria will resolve the difference regardless of sharing internal evidence. While these recommendations will benefit some variants, many variants identified in Mendelian disease testing are extremely rare¹ underscored by the observation that 88% of the 49,734 unique variants in this study were submitted by only one of the four laboratories. In addition, analysis of the sequencing data from 60,706 individuals in the Exome Aggregation Consortium (ExAC) found that 99% of identified high-quality variants have an allele frequency less than 1% and 54% of identified high-quality ExAC variants are only seen once in the entire dataset.¹² These findings further emphasize that the vast majority of variants identified by clinical testing laboratories are incredibly rare or even private and thus are unlikely to be regularly reassessed or potentially ever interpreted again. Therefore it is critical that laboratories share their observations and identify differences which may be the only prompt for reassessment.

Differences in interpretation of data are not unique to clinical genetic laboratories, as other healthcare providers are also subject to discordance in professional interpretation. When comparing pathologists' interpretations of individual breast biopsies to consensus-derived reference diagnosis, the overall concordance was only 75.3%.¹³ Similarly, radiologists were only 83% concordant in assessing dense versus nondense status of breast mammograms.¹⁴ These findings suggest that a consensus-based approach of multiple experts collaborating on a diagnosis may be necessary to improve interpretation. Similar approaches have been seen in rare disease genetics with the formation of international expert consortia that can provide consensus interpretations of variants, such as the International Society for Gastrointestinal Hereditary Tumours (InSiGHT) and Evidence-based Network for the Interpretation of Germline Mutation Alleles (ENIGMA).^{15,16} These disease-focused efforts not only apply expert interpretation but can organize data gathering efforts and spur research studies that can lead to a better understanding of genetic variation in their respective domains. For example, InSiGHT reclassified 54% of variants submitted with uncertain significance interpretations to pathogenic or likely pathogenic.¹⁵ To expand these efforts, ClinGen has formed numerous disease expert groups (<https://www.clinicalgenome.org/working-groups/clinical-domain/>) and continues to solicit further applications from groups to be considered as expert panels (three star level) with respect to their ClinVar submission status (http://www.ncbi.nlm.nih.gov/clinvar/docs/review_guidelines/). Once a three or four star variant interpretation is submitted to ClinVar from an experts, it will override submissions of lower review status to continually improve the reliability of public databases for use in diagnosis. However, these efforts heavily benefit from single lab submissions to identify variants that should be subject to expert review.

Laboratories did not reassess all 724 variants with interpretation differences as the goal of this pilot project was to determine if sharing internal data and implementing the ACMG-AMP guidelines could resolve difference and to setup a framework for addressing all differences in ClinVar. In continuing to reassess the remaining variants with interpretation differences between these four clinical laboratories, further emphasis will be placed on engaging additional ClinVar submitters. To promote more consistent variant interpretations in this field, participation from all submitters on a variant with conflicting interpretations is necessary to move more variants from a single-star review status to a two-star review status indicating that all submitted interpretations of a variant are concordant. This process will be overseen by ClinGen's Sequence Variant Inter-Laboratory Resolution working group, which will work on addressing all interpretation differences in ClinVar from all submitters. The lessons learned from this pilot project will guide the support and infrastructure needed to identify differences and facilitate sharing of internal data and classification rationale. For instance, in this study we found that 36% of interpretation differences were resolved simply by reassessing an older interpretation with laboratories' updated classification criteria. Given the time commitment to this project (an estimated 1–2 hours per variant per laboratory), future resolution will begin by encouraging laboratories with older interpretations or outlier interpretations to first reassess the variant with current guidelines, which would minimize the total time commitment for all laboratories.

In conclusion, these results demonstrate that clinical interpretations from these four clinical laboratories are now concordant for 91.5% of shared variants in ClinVar, with only 2% of

variant interpretation differences considered medically significant. Further specification regarding gene-specific criteria and expert guidance on the relevance of conflicting pathogenic and benign criteria may further facilitate resolution of interpretation differences with ACMG-AMP guidelines. Given that the interpretation of variants for their role in disease requires expert opinion and subjective review of scientific evidence and medical data, complete concordance is not expected. However, increased training and guidance on the application of the ACMG-AMP criteria and ensuring full sharing of evidence and classification rationales, is critical to move toward more consistent variant interpretations which will improve the care of patients with, or at risk for, genetic disorders.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Research reported in this publication was supported by the National Human Genome Research Institute (NHGRI), in conjunction with additional funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) under award number: U41HG006834. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We thank all ClinVar submitters for sharing their data and Jim Carver and Jessica Burgess for creating the figures.

References

1. Rehm HL, Berg JS, Brooks LD, et al. ClinGen--the Clinical Genome Resource. *N Engl J Med*. 2015; 372(23):2235–2242. [PubMed: 26014595]
2. Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016; 44(D1):D862–868. [PubMed: 26582918]
3. Amendola LM, Dorschner MO, Robertson PD, et al. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res*. 2015; 25(3):305–315. [PubMed: 25637381]
4. Yorczyk A, Robinson LS, Ross TS. Use of panel tests in place of single gene tests in the cancer genetics clinic. *Clin Genet*. 2015; 88(3):278–282. [PubMed: 25318351]
5. Pepin MG, Murray ML, Bailey S, et al. The challenge of comprehensive and consistent sequence variant interpretation between clinical laboratories. *Genet Med*. 2016; 18(1):20–24. [PubMed: 25834947]
6. Balmana J, Digiovanni L, Gaddam P, et al. Conflicting Interpretation of Genetic Variants and Cancer Risk by Commercial Laboratories as Assessed by the Prospective Registry of Multiplex Testing. *J Clin Oncol*. 2016
7. Amendola LM, Jarvik GP, Leo MC, et al. Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. *Am J Hum Genet*. 2016
8. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015; 17(5):405–424. [PubMed: 25741868]
9. Green RC, Goddard KA, Jarvik GP, et al. Clinical Sequencing Exploratory Research Consortium: Accelerating Evidence-Based Practice of Genomic Medicine. *Am J Hum Genet*. 2016; 98(6):1051–1066. [PubMed: 27181682]
10. Maxwell KN, Hart SN, Vijai J, et al. Evaluation of ACMG-Guideline-Based Variant Classification of Cancer Susceptibility and Non-Cancer-Associated Genes in Families Affected by Breast Cancer. *Am J Hum Genet*. 2016; 98(5):801–817. [PubMed: 27153395]

11. Jarvik GP, Browning BL. Consideration of Cosegregation in the Pathogenicity Classification of Genomic Variants. *Am J Hum Genet.* 2016; 98(6):1077–1081. [PubMed: 27236918]
12. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016; 536(7616):285–291. [PubMed: 27535533]
13. Elmore JG, Longton GM, Carney PA, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA.* 2015; 313(11):1122–1132. [PubMed: 25781441]
14. Sprague BL, Conant EF, Onega T, et al. Variation in Mammographic Breast Density Assessments Among Radiologists in Clinical Practice: A Multicenter Observational Study. *Ann Intern Med.* 2016
15. Thompson BA, Spurdle AB, Plazzer JP, et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet.* 2014; 46(2):107–115. [PubMed: 24362816]
16. Spurdle AB, Healey S, Devereau A, et al. ENIGMA--evidence-based network for the interpretation of germline mutant alleles: an international initiative to evaluate risk and clinical significance associated with sequence variation in BRCA1 and BRCA2 genes. *Hum Mutat.* 2012; 33(1):2–7. [PubMed: 21990146]

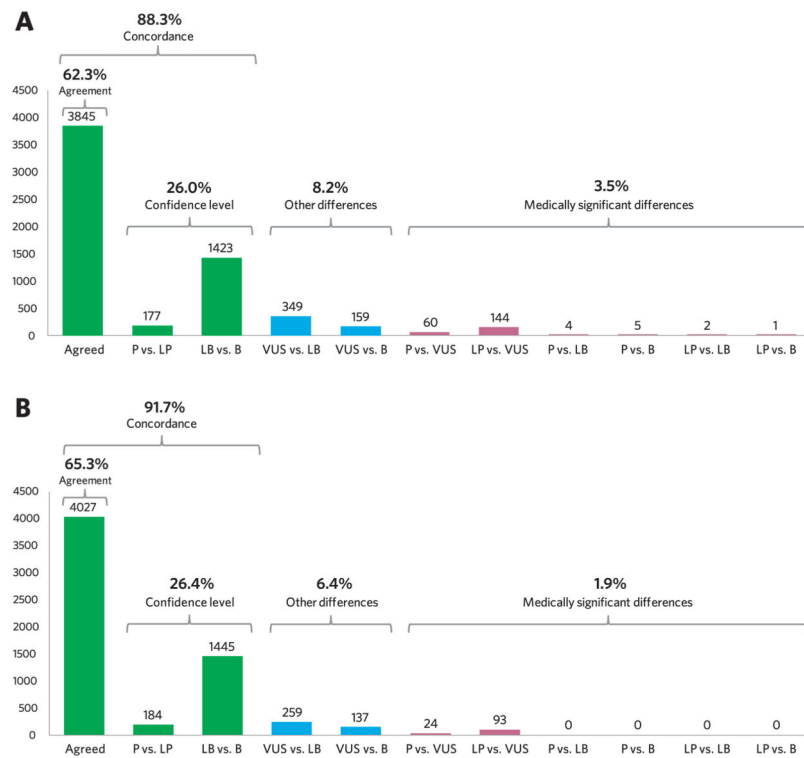


Figure 1. Distribution of variant interpretation comparisons between four clinical laboratories
 (A) Interpretation comparison of data in ClinVar (as of January 1, 2016), pre-resolution efforts
 (B) Interpretation comparison after reassessing 33% (242/724) of shared variants with interpretation differences

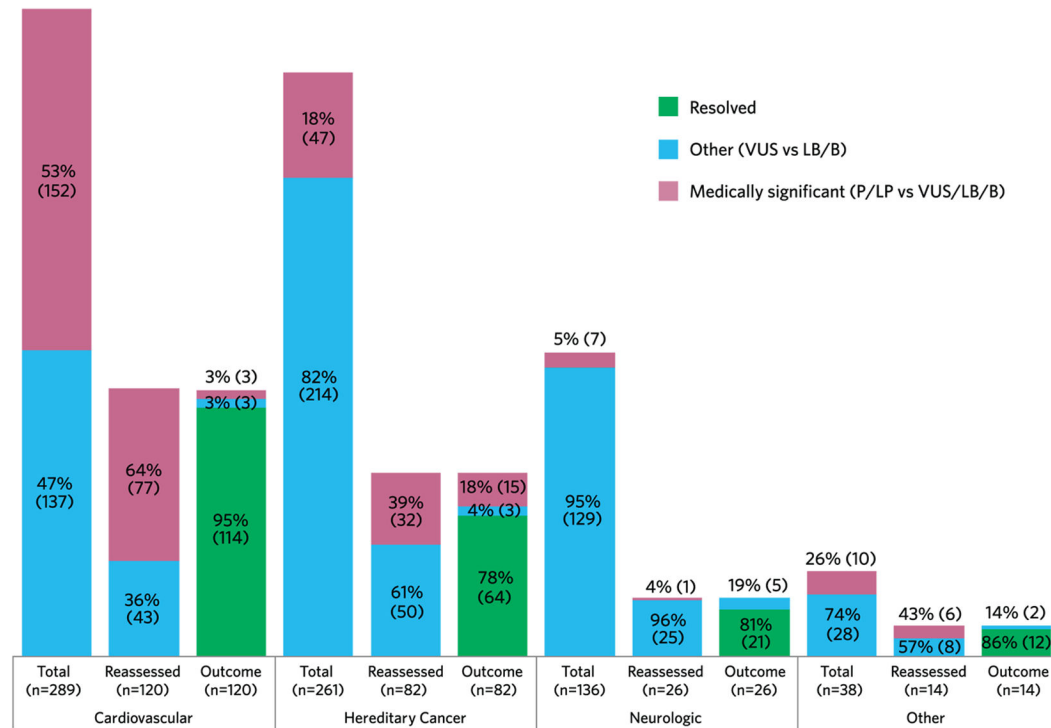


Figure 2. Distribution of interpretation differences and resolution outcome per disease areas
 Distribution of medically significant (P/LP vs VUS/LB/B) and other (VUS vs LB/B) differences within each disease area for all initial differences (“Total”), variants reassessed by laboratories (“Reassessed”), and final outcome for reassessed variants, including proportion resolved (“Outcome”).

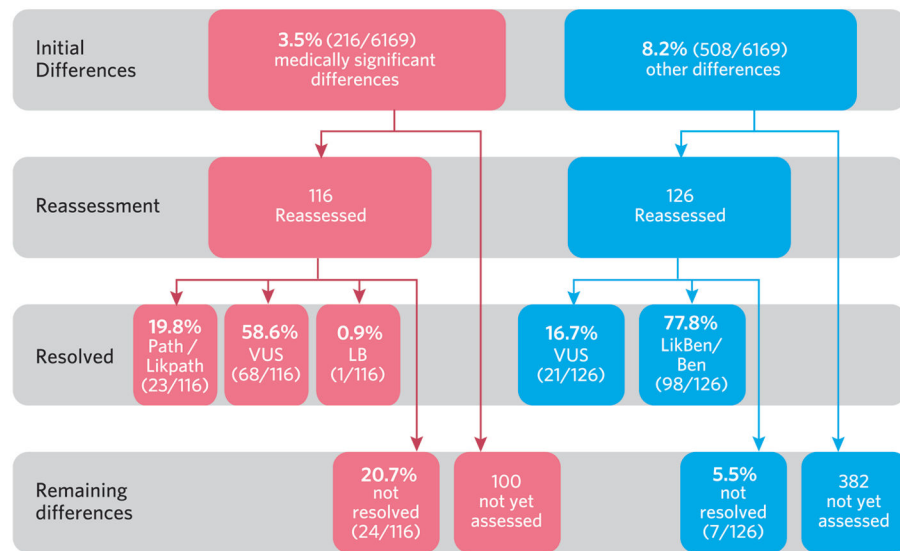


Figure 3.
Flowchart and outcome of variant resolution efforts.

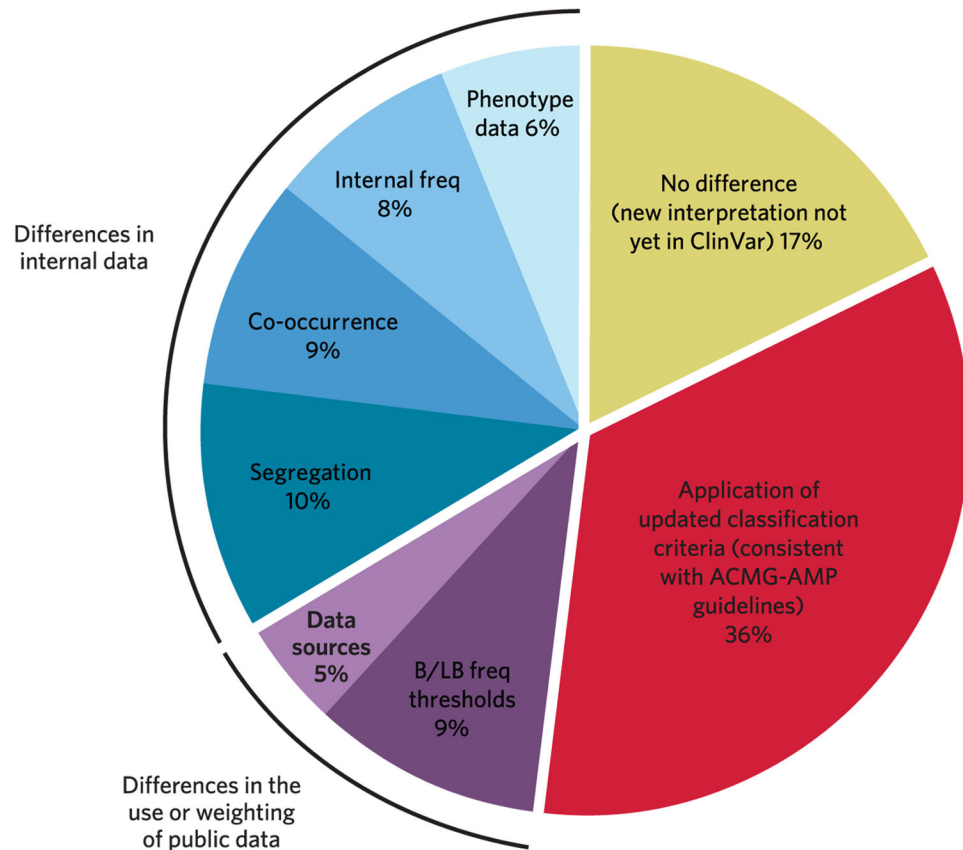


Figure 4. Basis of initial interpretation differences for resolved variants

Over half of the initial interpretation differences were resolved simply because the re-interpretation had already been completed but was not yet submitted to ClinVar (17%, yellow shading) or by reassessing an old variant interpretation with the laboratory's updated classification criteria, consistent with ACMG-AMP guidelines (36%, red shading). Differences in the use or weighting of public data accounted for 14% of interpretation differences (purple shading), including benign/likely benign thresholds (9%), and different data sources (5%). Differences in internal data accounted for 33% of interpretation differences (blue shading), including segregation data (10%), co-occurrence data (9%), internal proband frequency (8%), and detailed phenotype data (6%).

Table 1

Unique ACMG/AMP criteria applied contributing to persistent interpretation differences grouped into evidence categories.

ACMG-AMP Evidence category	% persistent differences impacted	ACMG-AMP Rule	Description	# P/LP vs. VUS/LB/B (%)	# VUS vs. LB/B (%)
Functional data	48%	PS3	Well-established functional studies show a deleterious effect	7 (29%)	0
		PM1	Mutational hot spot or well-studied function domain without benign variation	5 (21%)	0
		PP2	Missense in gene with low rate of benign missense variants and path missense common	1 (4%)	0
		BS3	Well-established functional studies show no deleterious effect	2 (8%)	0
Population data	45%	BS1	MAF is too high for disorder	4 (17%)	4 (57%)
		BS2	Observation in controls inconsistent with disease penetrance	3 (13%)	1 (14%)
		PS4	Prevalence in affected statistically increased over controls	2 (8%)	0
Other databases	29%	PP5	Reputable source = pathogenic	8 (33%)	0
		BP6	Reputable source = benign	1 (4%)	0
Computational & predictive data	16%	BP4	Multiple lines of computational evidence suggest no impact on gene product	1 (4%)	1 (14%)
		PM5	Novel missense change at a residue where a different path missense change has been seen	2 (8%)	0
		PP3	Multiple lines of computational evidence suggest a deleterious effect on gene product	1 (4%)	0
Other data	13%	PP4	Patient's phenotype or FH highly specific for gene	3 (13%)	0
		BP5	Found in case with an alternative cause	0	1 (14%)
Segregation data	6%	PP1	Co-segregation with disease in multiple affected family members	0	1 (14%)
		BS4	Non-segregation with disease	1 (4%)	0
Allelic data	3%	PM3	For recessive disorders, detected in trans with a pathogenic variant	1 (4%)	0

Note: ACMG-AMP benign criteria have codes beginning with "B" and pathogenic criteria have codes beginning with "P"